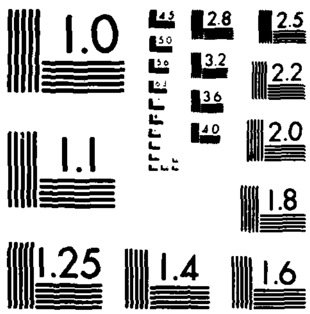
AD-A128 806 CONSTRUCTION OF NEW AREA SAMPLING FRAMES USING LANDSAT IMAGERY(U) ARMY ENGINEER TOPOGRAPHIC LABS FORT BELVOIR VA D J COSTANZO 17 MAR 83 1/1

UNCLASSIFIED F/G 8/2 NL

END
DATE FILMED
7-83
DTIC


1.0
2.8
2.5
3.2
2.2
3.6
4.0
2.0
1.1
1.8
1.25
1.4
1.6

CONSTRUCTION OF NEW AREA SAMPLING FRAMES USING LANDSAT IMAGERY

Daniel J. Costanzo*
U.S. Army Engineer Topographic Laboratories
Fort Belvoir, Virginia 22060

Date: 17 March 1983

BIOGRAPHICAL SKETCH

Mr. Costanzo is a Physical Scientist at the U.S. Army Engineer Topographic Laboratories. He transferred there from the U.S. Department of Agriculture (USDA), Statistical Reporting Service. While at the USDA, he helped develop several Area Sampling Frames, and assisted in teaching foreign agricultural experts about Area Sampling Frame construction. Mr. Costanzo holds a B.S. degree in Cartographic Science from The George Washington University, Washington, D.C. He is a member of ASP.

ABSTRACT

The U.S. Department of Agriculture develops Area Sampling Frames (ASF's) for each state using a series of land-use maps, and utilizes them in selecting statistical samples for agricultural surveys. Recently, Landsat imagery provided up-to-date land-use information during construction of a new ASF for California. Multispectral Scanner false-color scenes were manually interpreted along with aerial photos, maps, and ground data, to define land-use according to density of cultivation. Finally, these boundaries were digitized, forming a computerized ASF data base. Use of this ASF improved statistical efficiency with reduced sample size for California agricultural surveys.

INTRODUCTION

American agriculture is a dynamic system requiring constant maintenance of accurate and timely statistics on crop and livestock production, agricultural prices, and farm resources. The U.S. Department of Agriculture (USDA) Statistical Reporting Service uses a variety of sampling techniques to provide these basic agricultural data. One such method involves the utilization of Area Sampling Frames (ASF's) (Huddleston 1976, Houseman 1975). Separate ASF's are constructed for each state (Figure 1) by first dividing the state into counties, and then categorizing land areas within each county into relatively homogeneous units of broad land-use, called strata. To save both time and expense in sample selection, each stratum is further divided into smaller parcels called count units. The complete ASF consists of the entire collection of these count units covering a state's total land area, without any overlap or

*Formerly with the U.S. Department of Agriculture, Statistical Reporting Service, Washington, D.C. 20250

omission. To conduct a survey, samples of count units are first chosen using a probability proportional to the number of subsamples of segments of land, called sample units, within each count unit. These sample units are typically 259 hectares (ha) in size. Finally, a sample unit within each chosen count unit is randomly selected for coverage by ground personnel during the annual June Enumerative Survey of crop acreage and livestock.

ASF's traditionally have been constructed through manual interpretation of the latest available 1:63,360 scale black-and-white mosiacs (termed photo-index sheets) of low-altitude aerial photography, obtained from the USDA Agricultural Stabilization and Conservation Service (ASCS). Broad land-use strata would first be defined on these photo-index sheets, and later transferred to 1:126,720 scale paper county highway maps. Count units would then be delineated on these maps, and their enclosed land areas manually measured. The county highway maps would also serve as the base maps upon which the ASF would be permanently recorded. Throughout this process, it was critical that count unit and strata boundaries follow relatively permanent, linear features, such as roads, railways, and waterways, recognizable on the ground to an observer aided by maps and aerial photographs.

Unfortunately, this method of ASF construction has four major drawbacks (Hanuschak and Morrissey 1977): (1) large areas of the U.S. lack up-to-date aerial photographic coverage; (2) a number of photo-index sheets may be required to

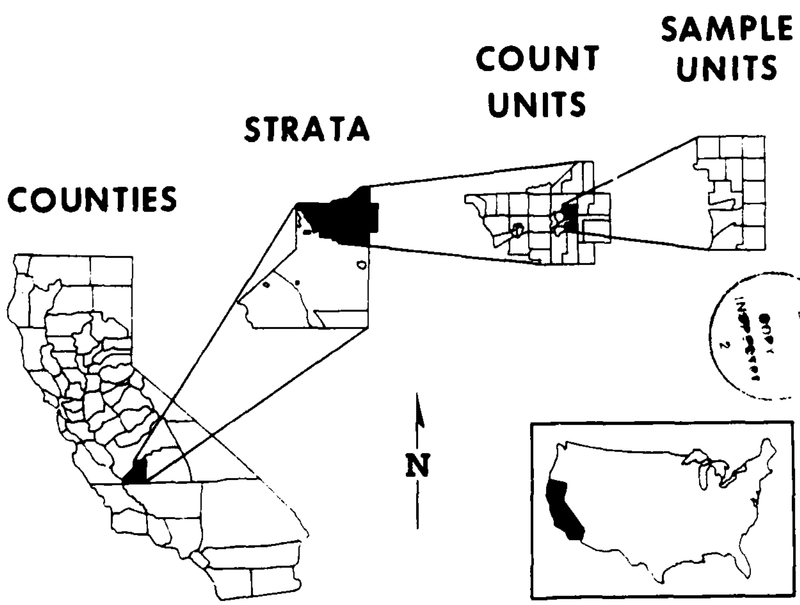


Figure 1. Hierarchy of components within an Area Sampling Frame (ASF). The region shown covers Kings County, California.

completely cover just one county, thus hindering an interpreter's ability to recognize broad land-use patterns; (3) the county highway maps are subject to loss or damage from frequent use; (4) the bulky paper format of the finished ASF impedes its use with new digital remote sensing techniques for crop area estimation (e.g., Bauer, et al. 1978). These digital techniques are dependent on ASF's in order to be properly undertaken (Hanuschak, et al. 1980; Wigton and Huddleston 1978). In 1978, the Statistical Reporting Service began employing Landsat Multispectral Scanner imagery in the ASF construction process, as well as digitization for computer storage of the completed ASF. This paper discribes how both of these new technologies are utilized for the first time, on an operational basis to construct a new ASF for California.

CONSTRUCTING THE CALIFORNIA ASF

California provided a good operational test for constructing a new ASF. It is a rapidly growing state significantly lacking complete, up-to-date ASCS aerial photographic coverage. Among it's 58 counties, eight contained no ASCS coverage, and ten posessed only incomplete coverage. Of the 40 remaining counties, only seven contained complete coverage dated 1971 or later. California also offered a diverse pattern of agricultural land-use characterized by large basic field sizes, which were well within the ground resolution of Landsat. And, the old California ASF, dating back to 1963, was due for replacement.

Unless extreme care is taken in the construction process, ensuing errors could produce substantial inaccuracies when the ASF is utilized for sample surveys (Wigton and Huddleston 1978). Therefore, the construction procedure (Figure 2) was separated into five sequential steps: input materials, initial stratification, final stratification, count unit installment, and digitization.

Input Materials

Thirty-five cloud-free Landsat-1 and 2 scenes dating from 13 May to 19 August 1976, were obtained to completely cover the state. All images were 1:250,000 false-color composite positive prints of Bands 4, 5, and 7. A scene index was also constructed by plotting the image borders on a 1:1,000,000 scale state map. This served as a ready reference for finding each scene's "footprint" on the ground. A mosaic of 1:1,000,000 scale black-and-white Band 5 prints of the scenes was also prepared. This mosiac became a valuable reference during analysis of gross land-use patterns over the entire state.

Even the oldest photo-index sheets could provide some useful information about boundaries or land-use, such as locating small roads in rural areas, or distinguishing orchards from general woodland. Therefore, the most recent 1:63,360 scale ASCS aerial photo-index sheets were obtained to supplement Landsat scene interpretation. The year of latest ASCS coverage ranged from 1952 to 1978. High altitude, 1978, U-2

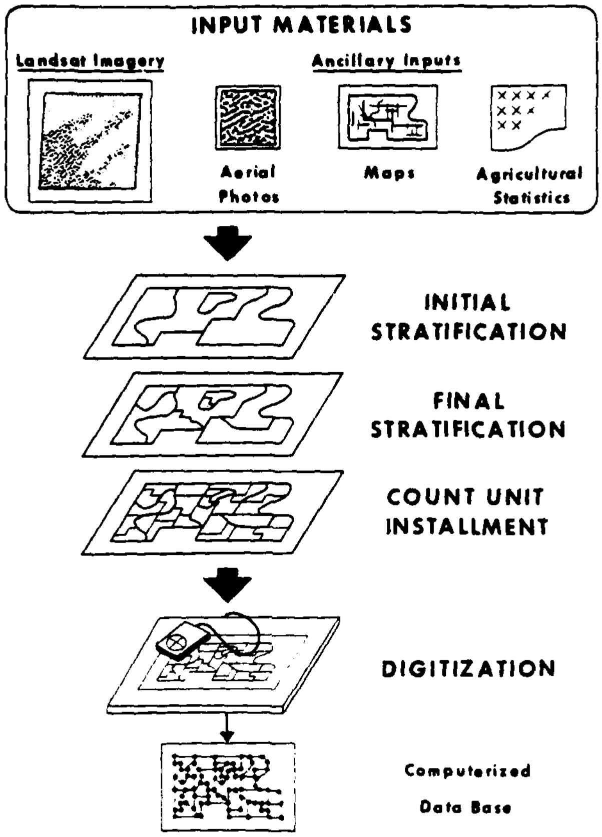


Figure 2. ASF construction procedure for each of the 58 counties in California.

aircraft black and white, 1:130,000 scale aerial photography was also obtained for limited areas of the state. They proved especially useful for urbanized areas, which, in general, were covered only by 1964 ASCS photography.

General county highway maps (1:126,720 scale) covering each county were also obtained, and photographically reduced and reproduced as transparencies matching the Landsat scene scale. They were used as overlays to the Landsat scenes during ASF construction. Supplemental 1:126,720 scale maps were also obtained from the U.S. Forest Service, the U.S. Department of the Interior Bureau of Land Management, and the California state government. U.S. Geological Survey 1:250,000, 1:62,500, and 1:24,000 scale topographic maps were also purchased to partially substitute for photo-index sheets in areas lacking photographic coverage.

It was quite important that agricultural land-use be known for a number of selected sites within each county. Data collected on 450 sample segments during the 1976 California June Enumerative Survey were used for this purpose. Each sample segment was identified on a computer printout summarizing agricultural land-use within, as determined through direct observation by ground personnel and interviews with farm operators. These data were obtained during the time period of Landsat coverage. Crop calendars and agricultural production reports for the state in general were supplied by the California Crop and Livestock Reporting Service.

This technique of ASF construction also required quantities of clear acetate for making overlays, as well as tape, marking pens, grease pencils, and colored pencils for marking and writing on the different photographic products, maps, and overlays. Simple hand lenses were the only optical equipment necessary.

Initial Stratification

Landsat imagery served at this stage of stratification to provide a broad regional overview showing general land-use trends within each county. Each false-color Landsat scene (Figure 3A) was registered to its respective county map overlay (Figure 3B) by lining up such prominent features as major roadways, rivers, lakes, reservoirs, and coastlines (the 1.6-by 1.6-km grid pattern present on Figure 3B represents the section line system common to the western U.S.). Local registration sufficient for visual interpretation was achieved by shifting the map overlay about the scene. Land-use was blocked off on an acetate overlay (Figure 3C) into six basic categories using only the Landsat scenes and county map overlays. These categories were identified by a single digit code number:

* Intensive Agriculture (1) - Areas 1,040 ha or larger exhibiting dense field patterns over 50 percent or more of their extent. Characterized by concentrated, geometric field shapes and association with water bodies. Colors ranged from white, pink, red, deep-red, blue, to black depending on crop growth status and field conditions.

* Extensive Agriculture (2) - Areas 1,040 ha or larger, displaying patches of concentrated cultivation over 15 to 50 percent of their extent. Characterized by the same shapes, associations, and colors as Intensive Agriculture, but with less dense field patterns. Usually located in fringe areas between intensive agriculture and range. Also consisted of cultivated areas in mountain valleys.

* Urban (3) - Built-up areas 65 ha or larger, identified by finely textured patterns of linear blue roadways, extending through blue-white commercial-industrial, and red-white residential areas. The ambiguous urban-rural boundaries on the Landsat scenes were sufficient for this preliminary stage of stratification.

* Range (4) - Regions 1,040 ha or larger, containing less than 15 percent cultivated land. Recognized by colors varying from red for woodland at higher elevations, to yellow-gray for dry forage vegetation, to tan-white for desert. Included public lands (belonging to the Forest Service or Bureau of Land Management) shown on the county map overlay.

* Non-Agriculture (5) - All land 1,040 ha or larger, not used for agricultural purposes and documented as such by law or other regulations. Including airports, wildlife refuges, national and state parks, and military installations shown on the county map overlay.

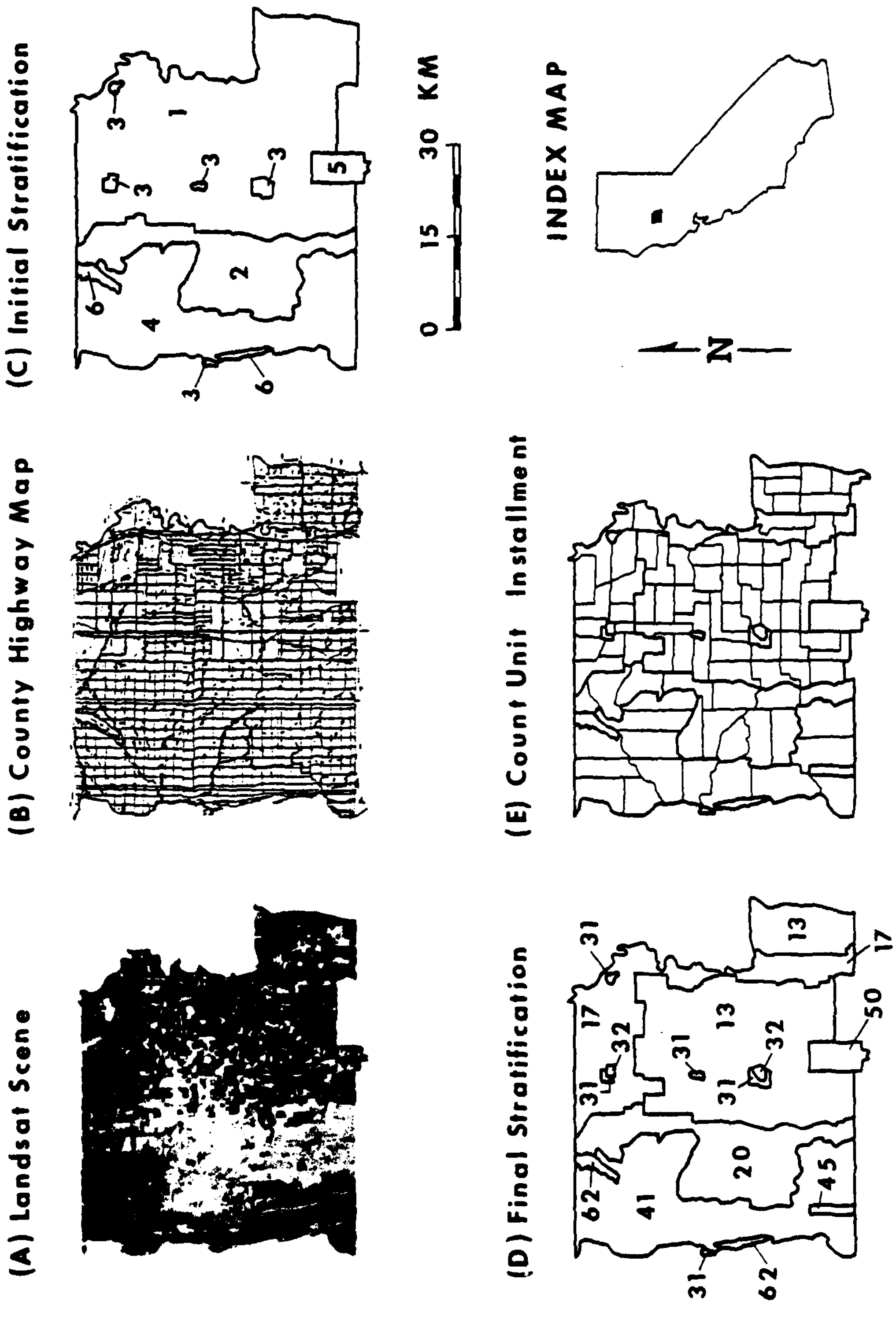


Figure 3. Progressive stages in ASF construction. The region shown is the eastern portion of Glenn County, California.

* Water (6) - Lakes and reservoirs 260 ha or larger, and major waterways greater than approximately 300 metres (m) in width. Characterized by their distinct blue-black color and shape. Shorelines shown on the county map overlay were used as boundaries for these water bodies to circumvent the problem of changing water levels.

Final Stratification

Upon completion of initial stratification for all counties, the ancillary information was examined to further refine what characteristics the six basic categories represented. This classification scheme, subdivided the six broad categories into twelve specific kinds of strata, each being identified by a double digit code number:

* General Crops (13) - Intensively cultivated land 520 ha or larger, containing mostly general crops with no more than 10 percent area coverage by fruit or vegetable crops. Characterized by large fields, up to 260 ha in size.

* Fruit Crops (17) - Intensively cultivated land 520 ha or larger containing primarily fruit crops mixed with general crops, and less than 33 percent vegetable crops. Recognized by field sizes much smaller than general crops, exhibiting a mottled reddish-white pattern, with tendency to cluster around urban areas.

* Vegetable Crops (19) - Intensively cultivated land 520 ha or larger not classified as Strata 13 or 17. Containing mostly vegetable mixed with general crops and very small amounts of fruit crops. Vegetable fields distinguishable from fruit orchards by being larger and more irregularly shaped.

* Extensive Cropland and Hay (20) - Same definition as Category 2 of initial stratification.

* Agri-Urban (31) - Areas 65 ha or larger containing agriculture mixed with more than 8 buildings per 100 ha. Did not include "strip" development along roadways in rural areas. Only aerial photography and large scale topographic maps would be used to identify Strata 31 and 32.

* Urban (32) - Heavily residential-commercial areas of major urban centers 65 ha or larger containing virtually no agriculture. More than 8 buildings per 100 ha.

* Private Range and Pasture (41) - Privately owned rangeland 1,040 ha or larger containing small, isolated areas of government land, less than 1,040 ha in size, not qualifying for other strata.

* Desert Range (43) - Barren areas 1,040 ha or larger containing very little agriculture of any kind.

* Public Grazing Allotments (44) - Public land 520 ha or larger containing livestock grazing allotments exceeding 520 ha. These allotments were identified on the supplemental maps.

* <u>Public Land with no known Agricultural Activity</u> (45) - Public land 1,040 ha or larger not containing grazing allotments, except for isolated parcels of private range and pasture too small (less than 1,040 ha in size) to be included separately under Stratum 41.

* <u>Non-Agriculture</u> (50) - Same definition as Category 5 of initial stratification.

* <u>Water</u> (62) - Same definition as Category 6 of initial stratification.

The image interpreters used these strata definitions to perform a final stratification of each county (Figure 3D), again on an acetate overlay over the Landsat scene. However, unlike for initial stratification, this final stratification required that the ancillary materials be employed with the Landsat scenes and county map overlays. Critical attention was paid to tying stratum boundaries to relatively permanent linear features recognizable to an observer on the ground using maps or aerial photographs. These fell into two major categories. The first was physical features: roads, railways, canals, and waterway borders. Field boundaries were not used because of the variability of cultivation patterns. Telephone, power, and pipelines were also avoided because of their tendency to cut across cultivated fields. The second category was non-physical features: national, state and county borderlines, and section lines. At this stage the interpreter relied almost exclusively on the aerial photographs, photo-index sheets, topographic maps, and county map overlays.

Count Unit Installment

Final strata boundaries were manually transferred from the acetate overlays onto paper county highway maps, and then subdivided into count units (Figure 3E). Because they formed the most permanent component of the ASF, count units had to be defined by relatively permanent boundaries, while maintaining a homogeneous land-use distribution within their enclosed land area. Count unit size also had to be such that a whole number of sample units could be assigned to each count unit. To permit this, Landsat imagery and ancillary data were relied on in the same manner as for final stratification, with a liberal tolerance being allowed around the average count unit size. Stratum 44, the public grazing land stratum was not broken down into count units. But for the purpose of simplicity, the term "count unit" in this paper refers to both true count units and Stratum 44 grazing allotments. The final step was to number these count units (and grazing allotments) for identification. When this process was completed, the paper county maps were ready for digitizing.

Digitization

The digitization operation utilized a high presision, manually-controlled, line following digitizer, directly interfaced, via modem hookup, to a minicomputer. This

allowed processing of the digitized data to be done in real time, using an interactive data analysis software system.

Digitization of each county highway map began with a calibration sequence to establish control points for converting map coordinates to geographic coordinates. The operator did this by digitizing at least four points common to a 1:250,000 scale topographic map and the county highway map. After calibration, count unit boundaries were digitized. The digitized information for each count unit was then entered on the minicomputer files in the following format: (1) count unit identification number; (2) calculated count unit area; and (3) geographic coordinates in latitude and longitude of digitized points marking the count unit border.

Upon completion of digitization, the digitized boundaries were outputed on a drum plotter at the same scale as the original county highway map. County map and output plot were then overlayed on a light table, thus providing an immediate error check. Any differences between the two were corrected by redigitizing the erroneous boundaries. The completed digitized ASF strata and count unit information was stored on a main frame computer at the USDA's Washington Computer Center. The new ASF was then put immediately into operational use for the first time to select 891 sample segments for sampling by ground personnel during the California June Enumerative Survey of 1979.

RESULTS

When the 1978 California June Enumerative Survey (sampled using the old ASF) and 1979 survey results were compared, it was found that the new ASF achieved better statistical efficiency than the old one (Fecso and Johnson 1981). Use of this new ASF to select a sample size one-third smaller than for 1978, provided reduced crop area sampling errors for three major crops (cotton, winter wheat, and rice) to coefficients of variation below the 12 percent precision limit considered usable in state survey estimates. The 1978 survey met this requirement for only one major crop category (cotton). In addition, coefficients of variation for pasture and non-agricultural area estimates were reduced from 14 percent in 1978 to 8 percent in 1979, and from 24 to 3 percent, respectively. The new ASF offers the ability to further reduce sampling errors to coefficients of variation as low as 6 percent for crop production estimates. Thus, area estimates for some crops which previously had unacceptable coefficients of variation now can be brought within usable levels. The major contributor to this reduction in sampling, as well as measurement (nonsampling) errors, is the improved delineation of land-use provided by the new ASF, particularly through utilization of current Landsat imagery.

After California, similar methods of ASF construction were employed to create new ASF's for the states of Washington, Oregon, Florida, Idaho, and Texas, plus the countries of Brazil and Argentina.

CONCLUSIONS

Manual interpretation of Landsat imagery offers an effective means for constructing new ASF's where there is a lack of comprehensive aerial photography. This kind of satellite imagery will become more important as the cost of aerial photography and ground-gathered agricultural information escalates. Further, digitization of the ASF puts it in a computer format compatible with these upcoming generations of agricultural remote sensing systems.

REFERENCES

Bauer, M. E.,M. M. Hixson, B. J. Davis, and J. B. Etheridge, 1978. "Area Estimation of Crops by Digital Analysis of Landsat Data," Photogrammetric Engineering and Remote Sensing, Vol. 44, No. 8, pp. 1033-1043.

Fecso, R., and V. Johnson, 1981. The New California Area Frame, U.S. Department of Agriculture, Statistical Reporting Service Report 22.

Hanuschak, G. A., and K. M. Morrissey, 1977. Pilot Study of the Potential Contributions of Landsat Data in the Construction of Area Sampling Frames, U.S. Department of Agriculture, Statistical Reporting Service.

Hanuschak, G. A., et al., 1980. "Crop-Area Estimates from Landsat; Transition from Research and Development to Timely Results," IEEE Transactions on Geoscience and Remote Sensing, Vol. GE-18, No. 2, pp. 160-166.

Houseman, E. E., 1975. Area Frame Sampling in Agriculture, U.S. Department of Agriculture, Statistical Reporting Service Report 20.

Huddleston, H. F., 1976. A Training Course in Sampling Concepts for Agricultural Surveys, U.S. Department of Agriculture, Statistical Reporting Service Report 21.

Wigton, W. H., and H. F. Huddleston, 1978. "A Land-Use Information System Based on Statistical Inference," Proceedings, Twelveth International Symposium on Remote Sensing of Environment, Environmental Research Institute of Michigan, Ann Arbor, Michigan, Vol. 1, pp. 429-442.